NÚMERO 520

VÍCTOR CARREÓN AND JORGE L. GARCÍA

Trivariate Probit with Double Sample Selection: Theory and Application

DICIEMBRE 2011

CIDE

www.cide.edu

Carretera México Toluca 3655, Col. Lomas de Santa Fe, 01210,
Álvaro Obregón, México DF, México.
www.cide.edu

Dirección de Publicaciones
publicaciones@cide.edu
Tel. 5081 4003

Abstract

We develop the trivariate probit model in which the sample incidentally truncates twice —i.e. in the first and in the second equations—, which is not solved in the literature. The model is analogue to the so called Bivariate Probit with Sample Selection (also referred as Bivariate Probit with Partial Partial Observabilty, Censored Probit or Heckman Probit) but in this case there are three equations and two truncations. We also present an application that shows the estimation biases when the incidental truncations are ignored.

Resumen

En este documento desarrollamos el modelo probit trivariado en el cual la muestra se trunca dos veces —en la primera y las segunda ecuaciones—, el cual no se resuelve en la literatura. El modelo es análogo al Probit con selección de muestra (también conocido como Probit bivariado con "partial partial Observabilty", Probit censurado o "Heckman Probit"); pero en nuestro caso hay tres ecuaciones y dos truncamientos. También presentamos una aplicación que muestra los sesgos de estimación cuando se omiten estos truncamientos.

1 Introduction

A sample selection bias refers to the situation whereby the individual observations studied are classified in a non-random way and, thus, divided into two or more subsamples. In practice, sample selection appears in two scenarios: 1) when the analyzed phenomenon is preceded by a natural selection process; 2) when the individual observations studied are self-selected. If sample selection is ignored selection bias –whose primer description and technical solution appears in Heckman (1979) –may arise.

The first scenario may be found, for example, when the determinants of the theft want to be known. If the reports are directly studied, ignoring the fact that they are preceded by a phenomenon such as theft, there may be a selection bias problem. It is therefore necessary to come up with a tool to determine whether theft and its reporting are statistically correlated and thus, verify whether it is correct to study the second phenomenon independently from the first.

In an econometric context, if the equations that describe theft and its reporting are statistically correlated, the independent estimation creates a selection bias problem. According to Sartori (2003) the bias appears because of two reasons: 1) individual observations that have higher propensity to suffer a theft are more likely to report a theft so one may observe a sample that has a non-random characteristic; 2) individual observations with low propensity to suffer a theft actually report it. This happens because they have high values on some unmeasured variables captured in the stochastic terms of the equation that characterizes theft. Hence, whether or not the independent variables in the theft reporting equation are uncorrelated with the stochastic term of the equation that characterizes theft in the overall population, the two variables are correlated in the selected sample. If the stochastic variables lead to a higher propensity to report heft, then we will have a bias in our estimation of the effect of the independent variables on it.

Therefore, we must estimate a bivariate model which considers a joint estimation of the two equations. This model should consider sample selection because in order to report a theft it is necessary to suffer it. The estimation must be found through the model illustrated in Diagram 1 if the dependent variables are dichotomous as it may be usual to suppose in this case.

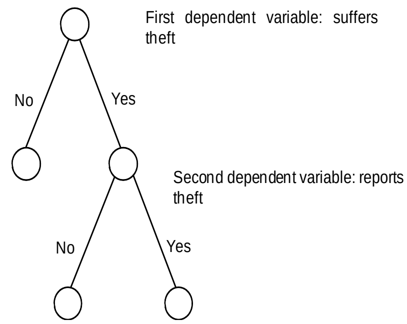

Diagram 1. Bivariate Probit with Sample Selection

The estimation of the model proposed in Diagram 1 makes it possible to find the correlation coefficient between the two dependent variables. Technically, this coefficient is useful because it shows whether the independent characterization of theft reporting is biased due to sample selection. If the correlation coefficient is statistically equal to zero, theft reporting can be characterized independently without any selection bias problems. If it is statistically different from zero, there will be a bias if the sample selection problem is ignored. Furthermore, since the direction of the bias is particular to each problem, it is important to estimate it.

The second scenario exemplifies attrition. In econometrics analysis, attrition is the non-random loss of information in a group of individual observations. For example, when some individuals respond to a survey incorrectly, the corresponding database offers no information on one or more variables.

If the attrition occurs in the dependent variable, the problem is known as truncation. Often, this problem is solved by proposing a model that considers truncated statistical distributions.[1] The non-random loss of information may also occur in the independent variables. For instance, it may happen that the individuals that constitute the studied sample decide not to provide information about socioeconomic variables. It is incorrect to assume, *a priori*, that this is a random decision. Therefore, the adequate answering of the survey must be considered as one of the phenomena that characterizes the individual observations.[2]

This situation may be illustrated in the previous situation. Let us suppose certain individual observations fail to adequately answer the survey from which theft and its reporting will be studied. It is then necessary to use a tool that studies the three

[1] This type of models are worked out in Greene (2008) p. 869-905.

[2] The adequate answering of the survey refers to the case where the relevant information that characterizes the equations to be estimated is actually observed by the researcher.

phenomena that take place jointly. This is because, in general, it is not true that the decision of answering the survey adequately or not is independent from, for example, the decision to report a crime or not. For example, it may happen that the individual observations that do not report theft are not interested in providing information regarding the crimes suffered, or that they fear reprisals for doing so. As explained before, this might carry a social cost, for example, the absence of information to prosecute crimes. Furthermore, it can create a positive correlation between reporting theft and answering the survey adequately.

Hence, estimating a pair of equations with dependent variables such as theft and reporting is not, *a priori*, enough to know the variables that impact these two actions. A third equation is needed to solve the attrition problem stemming from the collection of the data. Consequently, it is necessary to consider a trivariate model, since the correlations between the correct answering of the survey, the theft and the reporting may be statistically different from zero. The model discussed is illustrated in Diagram 2 and, as is the previous case, it may be natural to assume that the dependent variables are dichotomous. It is important to note that the order set forth in Diagram 2 is not the one in which the events took place, but that in which the data are presented. That is, when carrying out the study, the data for the variable theft are obtained if the survey is answered correctly. Subsequently, the data for the reporting are collected if a theft was suffered. Actually, this is the very reason why the double sample selection appears.

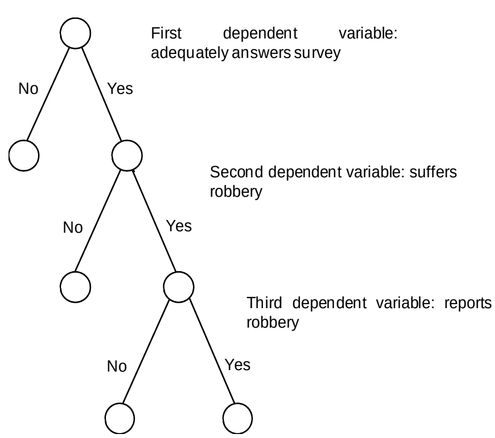

Diagram 2. Trivariate Probit Model with Double Sample Selection

As in the case illustrated in Diagram 1, the statistical magnitude of the correlation coefficients makes it possible to know what the correct characterization of the phenomena is. For example, if the correlations between the first and second, and the first and third dependent variables are statistically equal to zero, the estimation proposed in Diagram 1 is correct.

Often, as presumed in Diagrams 1 and 2, the variables that characterize theft and its reporting have observability limitations. For example, they are not characterized by continuous values, because they are dichotomous. That is, the variables take on an arbitrary value if the individual observations are victims of theft, and on another arbitrary value if they are not. In fact, each of the three dependent variables has the same characteristic : dichotomy. This makes the estimation set forth in Diagram 2 difficult to calculate and to interpret.

On one hand, several numerical integration methods must be combined in order to estimate the coefficients that characterize each one of the equations, and it is not always possible to find a solution. On the other hand, the characteristics of the dependent variable cause the *per se* estimation of the coefficients to have no interpretation, so that to know the impacts of the independent variables on the dependent variables, lineal transformations of the coefficients must be found. In practice, such transformations are usually complicated as well.

Therefore, facing the existence of situations like the one describe in Diagram 2, two realities are found in practice: 1) the sample selection issue caused by attrition is ignored, and the problem illustrated in Diagram 1 is directly tackled; 2) the database whose variables cause the attrition problem is disposed of, and a new database generated by adequately answered surveys is found. As previously explained, the first option is not theoretically robust. Thus, the results cannot be reliable in any context. The second option, on the other hand, is viable when information generation is abundant. However, it is not always possible to find information that simplifies the data analysis process. Then, the best way to tackle this kind of problem is to develop a *TPDSS*, which is not analyzed in the econometric literature.

There are two references with regard to the Trivariate Probit with Double Sample Selection (*TPDSS*) are found in the literature. On the one hand, Meng & Schmidt (1985) develop the Bivariate Probit Model with Sample Selection (*BPSS*) which corresponds, for example, to the proposition outlined in Diagram 1. On the other hand, Nielsen & Holm (2006) discuss the estimations of the Trvariate Heckit Model, which is analogue to the model developed by Heckman (1979) that jointly estimates a model that involves a dichotomous variable and a continuous variable. Also, they show some results on the estimations of a Trivariate Probit Model that has no sample selection problems. Cappellari & Jenkins (2003) give the generals of the Multivariate Probit estimations in Stata using simulated Maximum Likelihood Methods.

The main objective of our paper develop the *TPDSS* and to present an application of it. In order to do so, the reminder of the document is ordered as follows. In Section 2 we present the model, while Section 3 presents the application. Section 4 gives final comments.

2 Model

To carry out the estimation set forth in Diagram 2, a trivariate model in which the three dependent variables are dichotomous is developed: the *TPDSS*.

As in the case of the Probit, because the three dependent variables are dichotomous, they follow the same convention. If the occurrence they represent is positive, they are assigned a value equal to 1; if it is negative, they are assigned a value equal to 0. In this context, the first dependent variable must be equal to 1 to observe the second variable. Likewise, to observe the third variable, the second one must be equal to 1. That is, to observe whether the individual observation was victim of theft, it must answer the survey correctly (i.e. the first dependent variable has a positive occurrence). In a similar fashion, to observe whether the individual observation reports a theft, it has to suffer one (i.e. the second dependent variable has a positive occurrence).[3]

As in the case of the Probit models developed in Greene (2008, p. 770-96), we presume the existence of three latent variables characterized by a vector of regressors, $\mathbf{X}_{ij}$, and a stochastic term, ε_{ij}, for $j = 1, 2, 3$. Thus

$$\begin{aligned} y_{i1}^* &= \mathbf{X}_{i1}\boldsymbol{\beta}_1 + \varepsilon_{i1} \\ y_{i2}^* &= \mathbf{X}_{i2}\boldsymbol{\beta}_2 + \varepsilon_{i2} \\ y_{i3}^* &= \mathbf{X}_{i3}\boldsymbol{\beta}_3 + \varepsilon_{i3} \end{aligned} \tag{1}$$

where β_j is the vector of coefficients to be estimated in order to determine the different

[3] As explained before, the order in which the events takes place is not related to the way in which the data are presented.

impacts of the regressors on the dependent variable j, for $j = 1, 2, 3$. Note that $\mathbf{X}_{ij} = \begin{bmatrix} x_{i1j} & \cdots & x_{iKj} \end{bmatrix}$ is a vector of dimension $(1 \times K)$ and $\beta_j = \begin{bmatrix} \beta_{j1} \\ \vdots \\ \beta_{jK} \end{bmatrix}$ is a vector of dimension $(K \times 1)$.[4] Therefore, the subscript ijk denotes the individual, the regressor and the equation, respectively.

In this case, scalar and spatial identification problems do not disappear. If normalizations as those explained in Greene (2008, p. 770-96) are done, the dependent variables have the following mapping:

$$y_{i1} = \begin{cases} 1 & \text{if } \ y_{i1}^* > 0 \\ 0 & \text{if } \ y_{i1}^* \leq 0 \end{cases} \tag{2}$$

$$y_{i2} = \begin{cases} 1 & \text{if } \ y_{i1}^* > 0 \ , \ y_{i2}^* > 0 \\ 0 & \text{if } \ y_{i1}^* > 0 \ , \ y_{i2}^* \leq 0 \\ \text{not observed} & \text{if } \ y_{i1}^* \leq 0 \end{cases} \tag{3}$$

$$y_{i3} = \begin{cases} 1 & \text{if } \ y_{i1}^* > 0 \ , \ y_{i2}^* > 0 \ , \ y_{i3}^* > 0 \\ 0 & \text{if } \ y_{i1}^* > 0 \ , \ y_{i2}^* > 0 \ , \ y_{i3}^* \leq 0 \ . \\ \text{not observed} & \text{if } \ y_{i1}^* \leq 0 \ \text{ or } \ y_{i2}^* \leq 0 \end{cases} \tag{4}$$

To jointly estimate β_1, β_2 and β_3, the *Maximum Likelihood* (*ML*) method is used. To that end, the following assumption is made:

$$\varepsilon_i = \begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \varepsilon_{i3} \end{bmatrix} \sim_{i.i.d.} N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{bmatrix} \right).$$

Note that the assumption $Var(\varepsilon_{ij}|\mathbf{X}_{i1}, \mathbf{X}_{i2}, \mathbf{X}_{i3}) = 1$ for $j = 1, 2, 3$ solves the scalar identification problem. Now, to carry out the estimation by *ML*, the likelihood function $L(\cdot)$ is:

[4] In order to shorten notation, the analytical working of the problem is done for the case where the three equations have the same number of regressors. The results are immediately generalizable when this is not the case.

$$
\begin{aligned}
L\left(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\boldsymbol{\beta}_{3,}\rho_{12,}\rho_{23,}\rho_{13}|\cdot\right) &= \prod_{i=1}^{N}\Pr\left(y_{i1}^{*}>0,y_{i2}^{*}>0,y_{i3}^{*}>0|\mathbf{X}_{i1},\mathbf{X}_{i2},\mathbf{X}_{i3}\right)^{y_{i1}y_{i2}y_{i3}} \\
&\cdot\Pr\left(y_{i1}^{*}>0,y_{i2}>0,y_{i3}\leq 0|\mathbf{X}_{i1},\mathbf{X}_{i2},\mathbf{X}_{i3}\right)^{y_{i1}y_{i2}(1-y_{i3})} \\
&\cdot\Pr\left(y_{i1}^{*}>0,y_{i2}\leq 0|\mathbf{X}_{i1},\mathbf{X}_{i2},\mathbf{X}_{i3}\right)^{y_{i1}(1-y_{i2})} \\
&\cdot\Pr\left(y_{i1}^{*}\leq 0|\mathbf{X}_{i1},\mathbf{X}_{i2},\mathbf{X}_{i3}\right)^{1-y_{i1}} \qquad (5)
\end{aligned}
$$

where $i=1,\ldots,N$ are the observations, $X_h = \begin{bmatrix}\mathbf{X}_{1h}\\ \vdots \\ \mathbf{X}_{Nh}\end{bmatrix}$ for $h=1,2,3$. The subscript ik denotes the individuals and the regressor, respectively.

The argument for the construction of this likelihood function is analogous to the one presented in the case of the Probit model from Greene (2008, p. 770-96).

In this case, if the results presented in **Propositions 2**, **3** and **4** from **Appendix 1** are used:

$$
\begin{aligned}
\Pr\left(y_{i1}^{*}>0,y_{i2}^{*}>0,y_{i3}^{*}>0|\cdot\right) &= \Phi_3\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1,\mathbf{X}_{i2}\boldsymbol{\beta}_2,\mathbf{X}_{i3}\boldsymbol{\beta}_3;\rho_{12,}\rho_{23,}\rho_{13}\right) \qquad (6) \\
\Pr\left(y_{i1}^{*}>0,y_{i2}^{*}>0,y_{i3}^{*}\leq 0|\cdot\right) &= \Phi_3\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1,\mathbf{X}_{i2}\boldsymbol{\beta}_2,-\mathbf{X}_{i3}\boldsymbol{\beta}_3;\rho_{12,}-\rho_{23,}-\rho_{13}\right) \\
\Pr\left(y_{i1}^{*}>0,y_{i2}\leq 0|\cdot\right) &= \Phi_2\left(\mathbf{X}_{1}\boldsymbol{\beta}_1,-\mathbf{X}_{i2}\boldsymbol{\beta}_2;-\rho_{12}\right).
\end{aligned}
$$

And, as exposed in the case of the *Probit* model:

$$
\Pr\left(y_{i12}^{*}\leq 0|\mathbf{X}_{i1},\mathbf{X}_{i2},\mathbf{X}_{i3}\right)=\Phi\left(-\mathbf{X}_{i}\boldsymbol{\beta}\right). \qquad (7)
$$

Then, the likelihood function $L(\cdot)$ can then be written as:

$$
\begin{aligned}
L\left(\boldsymbol{\beta}_1,\boldsymbol{\beta}_2,\rho|\mathbf{X}_1,\mathbf{X}_2,\mathbf{X}_3\right) &= \prod_{i=1}^{N}\Phi_3\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1,\mathbf{X}_{i2}\boldsymbol{\beta}_2,\mathbf{X}_{i3}\boldsymbol{\beta}_3;\rho_{12,}\rho_{23,}\rho_{13}\right)^{y_{i1}y_{i2}y_{i3}} \qquad (8) \\
&\cdot\Phi_3\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1,\mathbf{X}_{i2}\boldsymbol{\beta}_2,-\mathbf{X}_{i3}\boldsymbol{\beta}_3;\rho_{12,}-\rho_{23,}-\rho_{13}\right)^{y_{i1}y_{i2}(1-y_{i3})} \\
&\cdot\Phi_2\left(\mathbf{X}_{1}\boldsymbol{\beta}_1,-\mathbf{X}_{i2}\boldsymbol{\beta}_2;-\rho_{12}\right)^{y_{i1}(1-y_{i2})} \\
&\cdot\Phi\left(-\mathbf{X}_{i}\boldsymbol{\beta}\right)^{1-y_{i1}}
\end{aligned}
$$

where $\phi(\cdot)$ is the p.d.f. of a standard normal random variable and $\Phi(\cdot)$ is its *c.d.f.*; $\phi_2(\cdot)$ is the *p.d.f.* of a bivariate standard normal random variable with means 0, variances 1 and correlation coefficient ρ_{12}, and its *c.d.f.* is $\Phi_2(\cdot)$; $\phi_3(\cdot)$ is the *p.d.f.* of a trivariate standard normal random variable with means 0, variances 1 and correlation coefficients ρ_{12}, ρ_{23}, ρ_{13} and $\Phi_3(\cdot)$ is its *c.d.f.*

After working out the estimation of the coefficients, the marginal effects are calculated to determine the impact of the regressors on the dependent variable. In this case, there are three relevant marginal effects: one on the expected value of the first dependent variable, $E(y_{i1}|\mathbf{X}_{i1}, \mathbf{X}_{i2}, \mathbf{X}_{i3})$, another on the expected value of the second dependent variable (taking into account that this is observed if and only if $y_{i1} = 1$), $E(y_{i2}|\mathbf{X}_{i1}, \mathbf{X}_{i2}, \mathbf{X}_{i3}, y_{i1} = 1)$, and the last one, on the expected value of the third dependent variable (knowing that this is observed if and only if $y_{i1} = y_{i2} = 1$), , $E(y_{i3}|\mathbf{X}_{i1}, \mathbf{X}_{i2}, \mathbf{X}_{i3}, y_{i1} = 1, y_{i2} = 1)$.

In the first case:

$$\begin{aligned} E(y_{i1}|\cdot) &= 0 \cdot \Pr\left(y_{i1}^* \leq 0|\mathbf{X}_{i1}, \mathbf{X}_{i2}, \mathbf{X}_{i3}\right) + 1 \cdot \Pr\left(y_{i1}^* > 0|\mathbf{X}_{i1}, \mathbf{X}_{i2}, \mathbf{X}_{i3}\right) \\ &= \Pr\left(\varepsilon_{i1} > -\mathbf{X}_{i1}\boldsymbol{\beta}_1|\mathbf{X}_{i1}, \mathbf{X}_{i2}, \mathbf{X}_{i3}\right) = \Pr\left(\varepsilon_{i1} < \mathbf{X}_{i1}\boldsymbol{\beta}_1|\mathbf{X}_{i1}, \mathbf{X}_{i2}, \mathbf{X}_{i3}\right) \\ &= \int_{-\infty}^{\mathbf{X}_{i1}\boldsymbol{\beta}_1} \phi(s)\, ds = \Phi\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1\right). \end{aligned} \tag{9}$$

Thus, if the regressor k is continuous, its marginal effect on the dependent variable is:

$$\begin{aligned} ME_{ik1} &= \frac{\partial E\left[y_{i1}|\mathbf{X}_{i1}, \mathbf{X}_{i2}, \mathbf{X}_{i3}\right]}{\partial x_{ik}} \\ &= \phi\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1\right)\beta_{1k} \end{aligned} \tag{10}$$

for $i = 1, \ldots, N$.

If the regressor k, is dichotomous, its marginal effect on the dependent variable is:

$$ME_{ik1} = \Phi\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1|x_{ik} = 1\right) - \Phi\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1|x_{ik} = 0\right) \tag{11}$$

for $i = 1, \ldots, N$.

The tests done on the coefficients are on the weighted average marginal effect. The *Delta Method* is then used to find out the distribution of the vector of marginal effects, as in the Probit and the *BPSS*.[5] Moreover, the distribution of the coefficients is analogue and the information matrix, $I(\boldsymbol{\beta})$, is calculated in the same way.[6]

In the second case:

$$\begin{aligned} E(y_{i2}|y_{i1} &= 1, \mathbf{X}_{i1}, \mathbf{X}_{i2}, \mathbf{X}_{i3}) = 0 \cdot \Pr\left(y_{i2}^* \leq 0 | y_{i1} = 1, \mathbf{X}_{i1}, \mathbf{X}_{i2}, \mathbf{X}_{i3}\right) \\ & +1 \cdot \Pr\left(y_{i2}^* > 0 | y_{i1} = 1, \mathbf{X}_{i1}, \mathbf{X}_{i2}, \mathbf{X}_{i3}\right) \\ &= \frac{\int\limits_{-\mathbf{X}_{i1}\boldsymbol{\beta}_1}^{\infty} \int\limits_{-\mathbf{X}_{i2}\boldsymbol{\beta}_2}^{\infty} \phi_2\left(s, t; \rho_{12}\right) ds dt}{\int\limits_{-\infty}^{\mathbf{X}_{i1}\boldsymbol{\beta}_1} \phi\left(t\right) dt} \\ &= \frac{\Phi_2\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1, \mathbf{X}_{i2}\boldsymbol{\beta}_2; \rho_{12}\right)}{\Phi\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1\right)}. \end{aligned} \tag{12}$$

As observed, in this case the result shown in **Proposition 1** of **Appendix 1** is used. If the regressor k is continuous and characterizes the two dependent variables,[7] its marginal effect on the second dependent variable is:

$$\begin{aligned} ME_{ik2} &= \frac{\partial E\left[y_{i2}|y_{i1} = 1, \mathbf{X}_{i1}, \mathbf{X}_{i2}, \mathbf{X}_{i3}\right]}{\partial x_{ik}} \\ &= \frac{\Phi\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1\right) \cdot \phi\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1\right) \cdot \Phi\left(\frac{\mathbf{X}_{i2}\boldsymbol{\beta}_2 - \rho_{12} \cdot \mathbf{X}_{i1}\boldsymbol{\beta}_1}{\sqrt{1-\rho_{12}^2}}\right) \cdot \beta_{1k}}{\Phi^2\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1\right)} \\ & + \frac{\Phi\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1\right) \cdot \phi\left(\mathbf{X}_{i2}\boldsymbol{\beta}_2\right) \cdot \Phi\left(\frac{\mathbf{X}_{i1}\boldsymbol{\beta}_1 - \rho_{12} \cdot \mathbf{X}_{i2}\boldsymbol{\beta}_2}{\sqrt{1-\rho_{12}^2}}\right) \cdot \beta_{2k}}{\Phi^2\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1\right)} \\ & - \frac{\Phi_2\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1, \mathbf{X}_{i2}\boldsymbol{\beta}_2; \rho_{12}\right) \cdot \phi\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1\right)}{\Phi^2\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1\right)} \cdot \beta_{1k} \end{aligned} \tag{13}$$

for $i = 1, \ldots, N$.

If the regressor k is dichotomous, its marginal effect on the dependent variable is:

[5] The *Delta Method* is explained in **Appendix 2.**

[6] In this case, ${}^t\boldsymbol{\beta} = \begin{bmatrix} \beta_{11} & \cdots & \beta_{1K} & \beta_{21} & \cdots & \beta_{2K} & \beta_{31} & \cdots & \beta_{3K} & \rho_{12} & \rho_{23} & \rho_{13} \end{bmatrix}$

[7] Note that if regressor k does not characterize equation l, then $\boldsymbol{\beta}_l = 0$ for $l = 1, 2$.

$$ME_{ik2} = \frac{\Phi_2\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1, \mathbf{X}_{i2}\boldsymbol{\beta}_2; \rho_{12}\right)}{\Phi\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1\right)}|_{x_{ik}=1} - \frac{\Phi_2\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1, \mathbf{X}_{i2}\boldsymbol{\beta}_2; \rho_{12}\right)}{\Phi\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1\right)}|_{x_{ik}=0} \tag{14}$$

for $i = 1, \ldots, N$.

The tests done on the coefficients are the same as in the former case. Also, the *Delta Method* is used to calculate the distribution of the vector of average marginal effects, which is similar to the one from the former case.

In the third case:

$$\begin{aligned}
E(y_{i3}|y_{i1} &= y_{i2} = 1, \mathbf{X}_{i1}, \mathbf{X}_{i2}, \mathbf{X}_{i3}) \qquad (15)\\
&= 0 \cdot \Pr\left(y_{i3}^{*} \leq 0 | y_{i1} = y_{i2} = 1, \mathbf{X}_{i1}, \mathbf{X}_{i2}, \mathbf{X}_{i3}\right) \\
&\quad +1 \cdot \Pr\left(y_{i3}^{*} > 0 | y_{i1} = y_{i2} = 1, \mathbf{X}_{i1}, \mathbf{X}_{i2}, \mathbf{X}_{i3}\right) \\
&= \frac{\int_{-\mathbf{X}_{i1}\boldsymbol{\beta}_1}^{\infty} \int_{-\mathbf{X}_{i2}\boldsymbol{\beta}_2}^{\infty} \int_{-\mathbf{X}_{i3}\boldsymbol{\beta}_3}^{\infty} \phi_3\left(s, t, r; \rho_{12}, \rho_{23}, \rho_{13}\right) dsdtdr}{\int_{-\infty}^{\mathbf{X}_{i1}\boldsymbol{\beta}_1} \int_{-\infty}^{\mathbf{X}_{i2}\boldsymbol{\beta}_2} \phi_2\left(s, t; \rho_{12}\right) dsd} \\
&= \frac{\Phi_3\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1, \mathbf{X}_{i2}\boldsymbol{\beta}_2, \mathbf{X}_{i3}\boldsymbol{\beta}_3; \rho_{12,}\rho_{23,}\rho_{13}\right)}{\Phi_2\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1, \mathbf{X}_{i2}\boldsymbol{\beta}_2; \rho_{12}\right)}.
\end{aligned}$$

As observed in this case, the result calculated in **Proposition 2** of **Appendix 1** is used. If the regressor k is continuous and characterizes the three dependent variables,[8] its marginal effect on the third dependent variable is:

[8]Note that if regressor k does not characterize equation j, then $\boldsymbol{\beta}_j = 0$ for $j = 1, 2, 3$.

$$
\begin{aligned}
ME_{ik3} &= \frac{\partial E\left[y_{i3}|y_{i1}=y_{i2}=1,\mathbf{X}_{i1},\mathbf{X}_{i2},\mathbf{X}_{i3}\right]}{\partial x_{ik}} \qquad (16)\\
&= \frac{1}{\Phi_2^2\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1,\mathbf{X}_{i2}\boldsymbol{\beta}_2;\rho_{12}\right)}\\
&\left(\begin{array}{l}
\left[\beta_{1k}\cdot\Phi_2\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1,\mathbf{X}_{i2}\boldsymbol{\beta}_2;\rho_{12}\right)\cdot\int\limits_{-\infty}^{\mathbf{X}_{i2}\boldsymbol{\beta}_2}\int\limits_{-\infty}^{\mathbf{X}_{i3}\boldsymbol{\beta}_3}\phi_3\left(\cdot\right)dsdt\right]\\
+\left[\beta_{2k}\cdot\Phi_2\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1,\mathbf{X}_{i2}\boldsymbol{\beta}_2;\rho_{12}\right)\cdot\int\limits_{-\infty}^{\mathbf{X}_{i1}\boldsymbol{\beta}_1}\int\limits_{-\infty}^{\mathbf{X}_{i3}\boldsymbol{\beta}_3}\phi_3\left(\cdot\right)dsdr\right]\\
+\left[\beta_{3k}\cdot\Phi_2\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1,\mathbf{X}_{i2}\boldsymbol{\beta}_2;\rho_{12}\right)\cdot\int\limits_{-\infty}^{\mathbf{X}_{i1}\boldsymbol{\beta}_1}\int\limits_{-\infty}^{\mathbf{X}_{i2}\boldsymbol{\beta}_2}\phi_3\left(\cdot\right)dtdr\right]\\
-\left[\beta_{1k}\cdot\Phi_3\left(\cdot\right)\cdot\phi\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1\right)\cdot\Phi\left(\frac{\mathbf{X}_{i2}\boldsymbol{\beta}_2-\rho_{12}\cdot\mathbf{X}_{i1}\boldsymbol{\beta}_1}{\sqrt{1-\rho_{12}^2}}\right)\right]\\
-\left[\beta_{2k}\cdot\Phi_3\left(\cdot\right)\cdot\phi\left(\mathbf{X}_{i2}\boldsymbol{\beta}_2\right)\cdot\Phi\left(\frac{\mathbf{X}_{i1}\boldsymbol{\beta}_1-\rho_{12}\cdot\mathbf{X}_{i2}\boldsymbol{\beta}_2}{\sqrt{1-\rho_{12}^2}}\right)\right]
\end{array}\right)
\end{aligned}
$$

for $i = 1, \ldots, N$.

If the regressor k, is dichotomous, its marginal effect on the third dependent variable is:

$$
\begin{aligned}
ME_{ik3} &= \frac{\Phi_3\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1,\mathbf{X}_{i2}\boldsymbol{\beta}_2,\mathbf{X}_{i3}\boldsymbol{\beta}_3;\rho_{12,}\rho_{23,}\rho_{13}\right)}{\Phi_2\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1,\mathbf{X}_{i2}\boldsymbol{\beta}_2;\rho_{12}\right)}\Big|_{x_{ik}=1} \qquad (17)\\
&-\frac{\Phi_3\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1,\mathbf{X}_{i2}\boldsymbol{\beta}_2,\mathbf{X}_{i3}\boldsymbol{\beta}_3;\rho_{12,}\rho_{23,}\rho_{13}\right)}{\Phi_2\left(\mathbf{X}_{i1}\boldsymbol{\beta}_1,\mathbf{X}_{i2}\boldsymbol{\beta}_2;\rho_{12}\right)}\Big|_{x_{ik}=0}
\end{aligned}
$$

for $i = 1, \ldots, N$.

The tests done on the coefficients are the same as in the former case. Also, the *Delta Method* is used to calculate the distribution of the vector of average marginal effects, which is similar to the one from the former case.

2.1 Statistical tests on the correlation coefficients

To examine the significance of the correlation coefficients, likelihood ratio tests are used. In this case, the null hypothesis is that the correlation coefficient(s) is/are equal to zero. According to Greene (2008, p. 820) :

$$-2\left(\left(\mathcal{L}\left(\theta_{R1}\right)+\cdots+\mathcal{L}\left(\theta_{Rn}\right)\right)-\mathcal{L}\left(\theta_{U}\right)\right)\sim\chi^{2}_{(q)}$$

where $\mathcal{L}\left(\theta_{Ri}\right)$ is the log-likelihood function for the restricted model evaluated at the $i = 1, \ldots N$ restricted estimates (i.e. carried out separately) and $\mathcal{L}\left(\theta_{U}\right))$ is the log-likelihood function for the unrestricted model evaluated at the unrestricted estimates (i.e. carried out jointly). Note that q is the number of degrees of freedom, equal to the number of coefficients whose value, under the null hypothesis, equals zero. Thus, for the estimation of the *BPSS* $q = 1$, whereas for the estimation of the trivariate Probit with sample selection $q = 3$.

3 Application

We use an example to illustrate our results and compare them with the standard results that will be obtained with the flawed approaches. Our data comes from the *Encuesta sobre victimización y eficacia institucional* (*Envei*) (*Survey on victimization and institutional efficacy*) conducted by the *Centro de Investigación y Docencia Económicas* (*CIDE*) (*Center for Research and Teaching in Economics*) carried on during 2007. The potential dependent variables of this database are summarized on **Table 1**.

Table 1. Distribution of the dependent variables

Adequate Answer of the Survey				
Yes		No		Total
# of observations	% of observations	# of observations	% of observations	# of observations
1080	72.68%	406	27.32%	1486
Adequate Answer of the Survey and Suffers a Theft				
Yes		No		Total
# of observations	% of observations	# of observations	% of observations	# of observations
347	32.13%	733	67.87%	1080
Adequate Answer of the Survey, Suffers a Theft, Reports Theft				
Yes		No		Total
# of observations	% of observations	# of observations	% of observations	# of observations
43	12.39%	304	87.61%	347

Source: Envei 2007.

The distribution of the second dependent variable, suffers theft, is conditional on the first dependent variable, answers the survey adequately. Similarly, the third dependent variable, reports the theft is conditional on the first and second dependent variables. It is important to remember that the conditionality of the dependent variables' distribution responds to the way the data are presented, not to the way in which the events take place. As observed, 27% of the households do not answer the survey correctly. It is then necessary to carry out an estimation that makes it possible to characterize the dependent variables comprised in the model illustrated in Diagram 2, so that theft and its reporting in Mexico City can be robustly described.

3.1 Data

The survey that spans our database was conducted during 2007 in the metropolitan area known as Mexico City -that is, the 16 boroughs of the Federal District and the 33 municipalities in the conurbation which belong to the State of Mexico. This is part of the *Programa de Seguridad Pública y Estado de Derecho* (*PESED*) (*Public Security and Rule of Law Program*) undertaken by the Law Faculty at *CIDE*. The sampling scheme used to carry out this survey was stratified in two stages. However, based upon certain population indices, weights were established for each unit of analysis, making it possible to treat the sample as representative of Mexico City's population.

Because of the way in which the *Envei 2007* presents the data, the units of analysis considered in this study are Mexico City's households. Consequently, all the data and results that appear in this study are at the household level. For example, in **Table 2** the variable "theft" shows whether any member of the household suffered at least one in 2007. The reasoning is similar for the remaining variables.

Table 2. Summary Statistics

Variables	Mean	s.d.
Awareness of authorities	.8189	.6964
Answers the survey correctly	.7268	.4458
Files a report	.1322	.3394
Education		
No schooling	.1373	.3443
Elementary	.1977	.4950
Secondary	.5385	.6844
Higher	.1306	.4012
Income		
income < 3 min. wages	.6530	.7241
3 min. wages < income < 10 min. wages	.3256	.7240
income > 10 min. wages	.0214	.1671
Locality		
State of Mexico	.4993	.6681
Federal District	.5007	.6681
Theft		
Auto Theft	.0115	.1145
burglary	.0266	.1884
Personal Theft	.1220	.3712
Household Size	2.8628	.7718
Average commute size		
time < 30 minutes	.2579	.6941
30 minutes < time < 90 minutes	.4369	.6545
time > 90 minutes	.3052	.5659
Type of transportation		
Public	.3158	.6042
Private	.6842	.6042

Source: Envei 2007.

The initial sample size was 1486 observations. The descriptive statistic of the variable "reporting" is conditional on theft taking place. Likewise, the descriptive statistic of "income" is conditional on the survey being answered correctly. The *Servicio de Administración Tributaria* (*Tax Administration Service*) (2010) reports, for 2007, a minimum wage of $50.57 pesos per day (in pesos of June 2002).

The variables that show the number of people and number of women per household are continuous. The variable "locality" is dichotomous: it equals 1 when the home is in the State of Mexico, and 0 when it is in the Federal District. The variable private transportation is dichotomous: equal to 1 when the average transportation of the household is private; equal to 0 when the average transportation is public. The variable "commute time" is dichotomous and has the three levels observed in **Table 2**, all representing the average time that the household's members spend transporting themselves. The variable "education" is dichotomous and has the three levels seen in **Table 2**; it refers to the head of the household's educational attainment. The variable "income" is treated in the same way as "education". The variables derived from the types of theft are dichotomous and equal to 1 if a member is a victim of any type of theft –note that they are exclusive and not in levels. Finally, note that the three dependent variables are dichotomous and equal to 1 if their occurrence is positive.

3.2 Results

Table 3 presents the results from the statistical test on the correlations coefficients of the *TPDSS*, explained in Section 2.1. As noted, we can reject the null hypothesis that the three correlation coefficients are statistically equal to zero with a significance of .01. This a sufficient condition to state that any estimation that ignores the attrition problem that appears in the first dependent variable and the selection problem that appears in the second dependent variable will lack of theoretical robustness.

Table 3. Joint Statistical Test on the Correlation Coefficients of the *TPDSS*

$\hat{\rho}_{12}$	$-.4436$
$\hat{\rho}_{13}$	.2231
$\hat{\rho}_{23}$	.0723
$\chi^2(3)$	8052.30
p. value	.0000

Table 4 presents various estimates. The first column shows the most naïve approach, where a model that considers theft reporting and ignores the other two dependent variables is estimated. The second column shows the estimation of the *BPSS*

that ignores attrition. The third column presents the estimation of the *TPDSS*. We just present the coefficients for the theft report equation to shortly exemplify the bias issue -of course, the reminder estimations are shown in **Appendix 3**. The difference between the estimations is considerable. However, we do not run any statistical test since the correlation test presented on **Table 3** is a sufficient condition to prove that the estimations on columns 1 and 2 are wrong and that the estimations on column 3 are theoretically robust. Note that this conclusion is not general and will depend on the specific dataset used to estimate the model. **Table 5** repeats the exercise of **Table 4** but for the case of the marginal effects.

Table 4. Coefficients Estimations for the Theft Report Equation

	[1]	[2]	[3]
Intercept	−2.5761	−3.5226	4.5932
	[1.4824]	[1.0926]	[.1688]
Locality (=1 if State of Mexico)	−.7001	−.5138	11.0253
	[.4271]	[.2981]	[.1860]
# People/Household	.1523	.0610	−.0337
	[.1642]	[.0964]	[200.4550]
% of Women/ Household	.0069	.0143	16.6769
	[.0114]	[.0085]	[1792.2403]
Awareness of authorities	−.2859	−.0011	17.1987
	[.5189]	[.3677]	[.3153]
Education			
Elementary	−.8889	−.6219	−21.4389
	[.8279]	[.5836]	[1.1420]
Secondary	.1877	.0957	−15.8431
	[.6907]	[.4675]	[.2747]
Higher	−.7896	−.2458	.6986
	[1.1670]	[.6467]	[1.1356]
Auto Theft	8.1372	3.4882	2.8355
	[315.3911]	[.5864]	[.4800]
Burglary	1.3144	1.2651	−4.0959
	[.6445]	[.4091]	[.3578]
Personal theft	.8601	1.8863	−10.0731
	[.5292]	[.3792]	[.3061]
Pseudo R^2	.8471	.8328	.5230
N	242	994	1486

Note: standard errors in brackets.

Table 5. Marginal Effects Estimations for the Theft Report Equation

	[1]	[2]	[3]
Locality (=1 if State of Mexico)	−.1451	−.0617	.0224
	[.0842]	[.0356]	[.0316]
# People/Household	.0315	.0069	.0365
	[.0326]	[.0112]	[07557]
% of Women/ Household	.0014	.0017	.0954
	[.0023]	[.0009]	[7.4971]
Awareness of authorities	−.0626	−.0001	.0321
	[.1190]	[.0439]	[.5731]
Education			
Elementary	−.1478	−.0585	.0000
	[.1038]	[.0609]	[.0021]
Secondary	.0378	.0126	−.0176
	[.1347]	[.0592]	[1.3729]
Higher	−.1257	−.0275	.0000
	[.1307]	[.0717]	[.0158]
Auto Theft	.7968	.8418	.1499
	[.0414]	[.0613]	[.0318]
Burglary	.3324	.2253	.1371
	[.1585]	[.0913]	[.0550]
Personal theft	.1573	.0924	−.0108
	[.0795]	[.0355]	[.0049]
Pseudo R^2	.8471	.8328	.5230
N	242	994	1486

Note: standard errors in brackets.

4 Final Comments

In this paper we analytically develop a Trivariate Probit Model with Double Sample Selection, an model that is absent in the econometric literature. In order to illustrate the relevance of the model, we present an application where the double selection happens with an attrition problem in the first equation and with a sample selection in the second equation. The results from the application show that, in this case, ignoring the aforementioned problems hinders the attainment of theoretically robust results.

References

[1] Ashford J. R. & Sowden R. R. (1970): Multivariate Probit Analysis, Biometrika, 26, 535-46.

[2] Cappelari L. & S. P. Jenkins (2003): Multivariate Probit Regression using Simulated Maximum Likelihood, Stata Journal, 3(3), 278-94.

[3] Greene, W. H. (2008): Econometric Analysis, 6th ed. Upper Saddle River, NJ: Prentice Hall.

[4] Heckman J.J. (1979): Selectivity Bias as a Specification Error, Econometrica, 47(1), 153-61.

[5] Meng C. & Schmidt P. (1985): On the Cost of Partial Observability in the Bivariate Probit Model, International Economic Review, 26(1), 71-85.

[6] Sartori A., (2003): An Estimator for Some Binary-Outcome Selection Models without Exclusion Restrictions, Political Analysis, 11, 111-38.

[7] Servicio de Administración Tributaria. (2007): Tipo de cambio (Exchange rate), Secretaria de Hacienda y Crédito Público, http://www.sat.gob.mx (accessed June 19, 2011).

[8] Servicio de Administración Tributaria. (2007): Salario mínimo por zona (Minimum wage by zone), Secretaria de Hacienda y Crédito Público, http://www.sat.gob.mx (accessed June 19, 2011).

Appendix 1

Proposition 1 *Let* $\phi_2(\cdot)$ *be the p.d.f. of a bivariate, random, standard, normal distributed with the means* 0*, variances* 1 *and correlation coefficient* ρ*. Also, consider*

$a, b, c \in \mathbb{R}$. *Then*

$$\int_{-a}^{c}\int_{-b}^{c} \phi_2(s, t; \rho)\, dsdt = \int_{-c}^{a}\int_{-c}^{b} \phi_2(s, t; \rho)\, dsdt. \tag{A1.1}$$

Proof. Define $s = -x$ and $t = -y$. Hence, $ds = -dx$ and $dt = -dy$. Then

$$\begin{aligned}\int_{-a}^{c}\int_{-b}^{c} \phi_2(s, t; \rho)\, dsdt &= -\int_{a}^{-c} -\int_{b}^{-c} \phi_2(-x, -y; \rho)\, dxdy \qquad (A1.2)\\ &= \int_{-c}^{a}\int_{-c}^{b} \phi_2(-x, -y; \rho)\, dxdy.\end{aligned}$$

Rewriting $\phi_2(\cdot)$,

$$\phi_2(-x, -y; \rho) = \phi_2(x, y; \rho) \tag{A1.3}$$

therefore

$$\begin{aligned}\int_{-c}^{a}\int_{-c}^{b} \phi_2(-x, -y; \rho)\, dxdy &= \int_{-c}^{a}\int_{-c}^{b} \phi_2(x, y; \rho)\, dxdy \qquad (A1.4)\\ &= \int_{-c}^{a}\int_{-c}^{b} \phi_2(s, t; \rho)\, dsdt\end{aligned}$$

which completes the proof. ■

Proposition 2 *Let $\phi_3(\cdot)$ be the p.d.f. of a trivariate, random, standard, normal distributed with the means 0, variances 1 and correlation coefficients ρ_{12}, ρ_{23}, ρ_{13}. Also, consider $a, b, c, d \in \mathbb{R}$. Then*

$$\int_{-a}^{d}\int_{-b}^{d}\int_{-c}^{d} \phi_3(s, t, r; \rho_{12}, \rho_{23}, \rho_{13})\, dsdtdr = \int_{-d}^{a}\int_{-d}^{b}\int_{-d}^{c} \phi_3(s, t, r; \rho_{12}, \rho_{23}, \rho_{13})\, dsdtdr. \tag{A1.5}$$

Proof. Define $s = -x$, $t = -y$ and $r = -z$. Hence $ds = -dx$, $dt = -dy$ and $dr = -dz$. Then

$$\int_{-a}^{d}\int_{-b}^{d}\int_{-c}^{d} \phi_3\left(s,t,r;\rho_{ij}\right) dsdtdr = \tag{A1.6}$$

$$-\int_{a}^{-d} -\int_{b}^{-d} -\int_{c}^{-d} \phi_3\left(-x,-y,-z;\rho_{ij}\right) dxdydz$$

$$= \int_{-d}^{a}\int_{-d}^{b}\int_{-d}^{c} \phi_3\left(-x,-y,-z;\rho_{ij}\right) dxdydz.$$

for $i,j = 1,2,3$ with $i \neq j$.

Rewriting $\phi_3(\cdot)$,

$$\phi_3(-x,-y,-z;\rho_{12},\rho_{23},\rho_{13}) = \phi_3(x,y,z;\rho_{12},\rho_{23},\rho_{13}) \tag{A1.7}$$

therefore

$$\int_{-d}^{a}\int_{-d}^{b}\int_{-d}^{c} \phi_3\left(-x,-y,-z;\rho_{ij}\right) dxdydz = \int_{-d}^{a}\int_{-d}^{b}\int_{-d}^{c} \phi_3\left(x,y,z;\rho_{ij}\right) dxdydz \tag{A1.8}$$

$$= \int_{-d}^{a}\int_{-d}^{b}\int_{-d}^{c} \phi_3\left(s,t,r;\rho_{ij}\right) dsdtdr$$

which completes the proof. ■

Proposition 3 *Let $\phi_3(\cdot)$ be the p.d.f. of a trivariate, random, standard, normal distributed with the means 0, variances 1 and correlation coefficients ρ_{12}, ρ_{23}, ρ_{13}. Also, consider $a,b,c,d \in \mathbb{R}$. Then*

$$\int_{-a}^{d}\int_{-b}^{d}\int_{-d}^{-c} \phi_3(s,t,r;\rho_{12},\rho_{23},\rho_{13})\, dsdtdr = \int_{-d}^{a}\int_{-d}^{b}\int_{-d}^{-c} \phi_3(s,t,r;\rho_{12},-\rho_{23},-\rho_{13})\, dsdtdr. \tag{A1.9}$$

Proof. Define $t = -y$ and $r = -z$. Hence, $dt = -dy$ and $dr = -dz$. Then

$$\int_{-a}^{d}\int_{-b}^{d}\int_{-d}^{-c} \phi_3\left(s,t,r;\rho_{ij}\right) dsdtdr = \tag{A1.10}$$

$$-\int_{a}^{-d} - \int_{b}^{-d}\int_{-d}^{-c} \phi_3\left(s,-y,-z;\rho_{ij}\right) dsdydz$$

$$= \int_{-d}^{a}\int_{-d}^{b}\int_{-d}^{-c} \phi_3\left(s,-y,-z;\rho_{ij}\right) dsdydz.$$

for $i,j = 1,2,3$ with $i \neq j$.

Rewriting $\phi_3\left(s,-y,-z;\rho_{12},\rho_{23},\rho_{13}\right)$,

$$\phi_3\left(s,-y,-z;\rho_{12},\rho_{23},\rho_{13}\right) = \phi_3\left(x,y,r;\rho_{12},-\rho_{23},-\rho_{13}\right) \tag{A1.11}$$

therefore

$$\int_{-d}^{a}\int_{-d}^{b}\int_{-d}^{-c} \phi_3\left(s,-y,-z;\rho_{ij}\right) dxdydr = \tag{A1.12}$$

$$\int_{-d}^{a}\int_{-d}^{b}\int_{-d}^{-c} \phi_3\left(s,y,z;\rho_{12},-\rho_{23},-\rho_{13}\right) dsdydz$$

$$= \int_{-d}^{a}\int_{-d}^{b}\int_{-d}^{-c} \phi_3\left(s,t,r;\rho_{12},-\rho_{23},-\rho_{13}\right) dsdtdr$$

which completes the proof. ■

Proposition 4 *Let $\phi_2(\cdot)$ be the p.d.f. of a bivariate, random, standard, normal distributed with the means 0, variances 1 and correlation coefficients ρ. Also, consider $a,b,c \in \mathbb{R}$. Then*

$$\int_{-a}^{d}\int_{-d}^{-c} \phi_2\left(s,t;\rho\right) dsdt = \int_{-d}^{a}\int_{-d}^{-c} \phi_2\left(s,t;-\rho\right) dsdt \tag{A1.13}$$

Proof. Define $t = -y$. Hence, $dt = -dy$. Then

$$\int_{-a}^{d}\int_{-d}^{-c} \phi_2(s,t;\rho)\,dsdt = -\int_{a}^{-d}\int_{-d}^{-c} \phi_2(s,-y;\rho)\,dsdy = \int_{-d}^{a}\int_{-d}^{-c} \phi_2(s,-y;\rho)\,dsdy. \quad (A1.14)$$

Rewriting $\phi_2(\cdot)$,

$$\phi_2(s,-y;\rho) = \phi(s,y;-\rho) \quad (A1.15)$$

therefore

$$\int_{-d}^{a}\int_{-d}^{-c} \phi_2(s,-y;\rho)\,dsdy = \int_{-d}^{a}\int_{-d}^{-c} \phi_2(s,y;-\rho) = \int_{-d}^{a}\int_{-d}^{-c} \phi_2(s,t;-\rho)\,dsdt \quad (A1.16)$$

which completes the proof. ■

Appendix 2

Delta Method

Let $\boldsymbol{\xi}$ be a $p \times 1$ vector. If $\boldsymbol{\xi} \sim N(\boldsymbol{\mu}, \Lambda)$ and $f(\boldsymbol{\xi})$ is any given function, then:

$$f(\boldsymbol{\xi}) \sim N\left(f(\boldsymbol{\mu}), J(f(\boldsymbol{\xi})) \cdot \Lambda \cdot^{t} J(f(\boldsymbol{\xi}))\right)$$

where $J(f(\boldsymbol{\xi}))$ is the Jacobian matrix of the function $f(\boldsymbol{\xi})$. Here, $\boldsymbol{\mu}$ is of dimension $p \times 1$ and Λ is of dimension $p \times p$.[9]

[9] This method is formally stated in Greene (2008, p. 1055).

Appendix 3

Table A1. Coefficients Estimations for the Attrition Equation

	Univariate, Bivariate, Trivariate
Intercept	.180
	[.182]
Locality (=1 if State of Mexico)	.060
	[.091]
Education	
Elementary	.004
	[.031]
Secondary	−.001
	[.002]
Higher	−.116
	[.098]
Auto Theft	−.413
	[.114]
Burglary	−.391
	[.120]
Personal theft	102.521
	[56.859]

Note: standard errors in brackets.

Table A2. Marginal Effects Estimations for the Attrition Equation

	Univariate, Bivariate, Trivariate
Locality (=1 if State of Mexico)	.021
	[.031]
Education	
Elementary	.001
	[.011]
Secondary	.000
	[.001]
Higher	−.040
	[.034]
Auto Theft	−.143
	[.039]
Burglary	−.136
	[.041]
Personal theft	.429
	[.064]

Note: standard errors in brackets.

Table A3. Coefficients Estimations for the Theft Equation

	Univariate	Bivariate	Trivariate
Intercept	.144	−.662	23.760
	[.364]	[.206]	[1406.712]
Locality (=1 if State of Mexico)	.055	.017	−.10.742
	[.161]	[.091]	[.353]
# People/Household	.115	.055	7.189
	[.065]	[.032]	[.128]
% of Women/ Household	−.009	−.006	1.050
	[.004]	[.002]	[.044]
Private Transportation	.022	.016	−1.640
	[.179]	[.105]	[.003]
Commute Time			
30 minutes < time < 90 minutes	.426	.266	−8.499
	[.194]	[.117]	[.163]
time > 90 minutes	.496	.283	−9.555
	[.206]	[.122]	[.232]
Income			
< 3 minimum wages	−.351	−.246	−2.898
	[.180]	[.096]	[.207]
> 10 minimum wages	−.093	−.055	3.463
	[.605]	[.307]	[.245]
Pseudo R^2	.400	.833	.523
N	994	994	1486

Note: standard errors in brackets.

Table A4. Marginal Effects Estimations for the Theft Equation

	Univariate	Bivariate	Trivariate
Locality (=1 if State of Mexico)	.020	.005	−.036
	[.059]	[.027]	[.216]
# People/Household	.042	.016	.024
	[.023]	[.009]	[.760]
% of Women/ Household	−.003	−.002	.003
	[.001]	[.001]	[.111]
Private Transportation	.008	−.005	−.556
	[.066]	[.031]	[.272]
Commute Time			
30 minutes < time < 90 minutes	[.155]	[.074]	[−.023]
	.068	.032	.272
time > 90 minutes	[.179]	[.079]	[−.032]
	.070	.033	.272
Income			
< 3 minimum wages	−.130	−.074	−.010
	[.066]	[.030]	[.275]
> 10 minimum wages	−.034	−.018	.012
	[.022]	[.097]	[.275]
Pseudo R^2	.400	.833	.523
N	994	994	1486

Note: standard errors in brackets.

Novedades

División de Administración Pública

Ma. Amparo Casar, *Representation and Decision Making in the Mexican Congress*, DTAP-258

Guillermo Cejudo y Dirk Zavala, *Procesos y prácticas del acceso a la información: un ejercicio de usuario simulado*, DTAP-257

Ernesto Flores-Roux y Judith Mariscal, *The Development of Mobile Money Systems*, DTAP-256

David Arellano *et al.*, *Control de los conflictos de interés*, DTAP-255

David Arellano, Víctor Figueras y Walter Lepore, *Política de tránsito en el DF de México: una introducción a su estudio empírico*, DTAP-254

Sergio Cárdenas y Maximiliano Cárdenas, *La participación del poder legislativo en la definición de la política educativa en México*, DTAP-253

Sergio Cárdenas, *Administración centrada en la escuela,* DTAP-252

Joanna D. Lucio, Edgar Ramírez y Sergio Cárdenas, *¿Libertad para quién? El efecto de comunidades cerradas en el espacio urbano*, DTAP-251

Edgar E. Ramírez, *Land Development Permitting*, DTAP-250

Rodrigo Sandoval-Almazán, Luis F. Luna-Reyes y J. Ramón Gil-García, *Índice de Gobierno Electrónico Estatal: La medición 2009*, DTAP-249

División de Economía

Rafael Garduño-Rivera and Esteban Fernández-Vázquez, *Ecological Inference with Entropy Econometrics: Using the Mexican Census as a Benchmark*, DTE-511

Rafael Garduño, Kathy Baylis and Mary P. Arends-Kuenning, *Regional Economic Analysis of Internal Migration in Mexico*, DTE-510

Brasil Acosta Peña and Andrés Aradillas López, *Semi-Empirical Likelihood Estimation of Manufacturing Interaction-Based Model…*, DTE-509

Brasil Acosta Peña, *Technical Efficiency in the Mexican Manufacturing Sector: A Stochastic Frontier Approach*, DTE-508

Brasil Acosta Peña, *Estimation of a Multiple Equilibrium Game with Complete Information…*, DTE-507

Juan Rosellón, Anne Neumann and Hannes Weigt, *Removing Cross-Border Capacity Bottlenecks in the European Natural Gas Market…*, DTE-506

Alejandro López and Javier Parada, *Poverty Dynamics in Rural Mexico: An Analysis Using Four Generations of Poverty Measurement*, DTE-505

Kaniska Dam and Antonio Jiménez-Martínez, *A Note on Bargaining over Complementary Pieces of Information in Networks*, DTE-504

David Mayer y Grodecz Ramírez, *Ciclo de vida humano y ciclo de vida urbano: Urbanización y desarrollo económico*, DTE-503

Kaniska Dam y Daniel Ruiz Pérez, *On the Existence of Sharecropping*, DTE-502

DIVISIÓN DE ESTUDIOS INTERNACIONALES

Rafael Velázquez, *La política exterior de Estados Unidos hacia México bajo la administración de Barack Obama*, DTEI-220

Rafael Velázquez y Jorge Schiavon, *La cooperación internacional descentralizada para el desarrollo en México*, DTEI-219

Kimberly A. Nolan García, *Whose Preferences?: Latin American Trade Promotion Pacts as a Tool of US Foreign Policy*, DTEI-218

Farid Kahhat, Rodrigo Morales y Ana Paula Peñalva, *El Perú, las Américas y el Mundo. Opinión pública y política exterior en Perú, 2010*, DTEI-217

Guadalupe González, Jorge Schiavon, David Crow y Gerardo Maldonado, *Mexico, the Americas and the World 2010…*, DTEI-216

Guadalupe González, Jorge Schiavon, David Crow y Gerardo Maldonado, *México, las Américas y el Mundo 2010. Política exterior: Opinión pública y líderes*, DTEI-215

Gerardo Maldonado, *Desapego político y desafección institucional en México: ¿Desafíos para la calidad de la democracia?*, DTEI-214

Luz María de la Mora, *Opening Markets without Reciprocity: Explaining Mexico's Trade Policy, 2000-2010*, DTEI-213

Farid Kahhat, *Las industrias extractivas y sus implicaciones políticas y económicas*, DTEI-212

Mariana Magaldi de Sousa, *Trade Openness and the Channels of its Impact on Democracy*, DTEI-211

DIVISIÓN DE ESTUDIOS JURÍDICOS

María Mercedes Albornoz , *Cooperación interamericana en materia de restitución de menores*, DTEJ-56

Marcelo Bergman, *Crimen y desempleo en México: ¿Una correlación espuria?*, DTEJ-55

Jimena Moreno, Xiao Recio y Cynthia Michel, *La conservación del acuario del mundo. Alternativas y recomendaciones para el Golfo de California*, DTEJ-54

María Solange Maqueo, *Mecanismos de tutela de los derechos de los beneficiarios*, DTEJ-53

Rodolfo Sarsfield, *The Mordida´s Game. How institutions incentive corruption*, DTEJ-52

Ángela Guerrero, Alejandro Madrazo, José Cruz y Tania Ramírez, *Identificación de las estrategias de la industria tabacalera en México*, DTEJ-51

Estefanía Vela, *Current Abortion Regulation in Mexico*, DTEJ-50

Adriana García and Alejandro Tello, *Salaries, Appelate Jurisdiction and Judges Performance*, DTEJ-49

Ana Elena Fierro and Adriana García, *Design Matters: The Case of Mexican Administrative Courts*, DTEJ-48

Gustavo Fondevila, *Estudio de percepción de magistrados del servicio de administración de justicia familiar en el Distrito Federal*, DTEJ-47

DIVISIÓN DE ESTUDIOS POLÍTICOS

Ana Carolina Garriga, *Regulatory Lags, Liberalization, and Vulnerability to Systemic Banking Crises*, DTEP-232

Rosario Aguilar, *The Tones of Democratic Challenges: Skin Color and Race in Mexico*, DTEP-231

Rosario Aguilar, *Social and Political Consequences of Stereotypes Related to Racial Phenotypes in Mexico*, DTEP-230

Raúl C. González and Caitlin Milazzo, *An Argument for the 'Best Loser' Principle in Mexico*, DTEP-229

Francisco Javier Aparicio and Covadonga Meseguer, *Supply or Demand? Politics and the 3x1 Program for Migrants*, DTEP-228

Ana Carolina Garriga and Brian J. Phillips, *Foreign Aid and Investment in Post-Conflict Countries*, DTEP-227

Allyson Benton, *The Origins of Mexico's Municipal Usos y Costumbres Regimes*, DTEP-226

Ana Carolina Garriga, *Objetivos, instrumentos y resultados de política monetaria. México 1980-2010*, DTEP-225

Andreas Schedler, *The Limits to Bureaucratic Measurement. Observation and Judgment in Comparative Political Data Development*, DTEP-224

Andrea Pozas and Julio Ríos, *Constituted Powers in Constitution-Making Processes. Supreme Court Judges, Constitutional Reform and the Design of Judicial Councils*, DTEP-223

DIVISIÓN DE HISTORIA

Michael Sauter, *Human Space: The Rise of Euclidism and the Construction of an Early-Modern World, 1400-1800*, DTH-75

Michael Sauter, *Strangers to the World: Astronomy and the Birth of Anthropology in the Eighteenth Century*, DTH-74

Jean Meyer, *Una revista curial antisemita en el siglo XIX: Civiltá Cattolica*, DTH-73

Jean Meyer, *Dos siglos, dos naciones: México y Francia, 1810- 2010*, DTH-72

Adriana Luna, *La era legislativa en Nápoles: De soberanías y tradiciones*, DTH-71

Adriana Luna, *El surgimiento de la Escuela de Economía Política Napolitana*, DTH-70

Pablo Mijangos, *La historiografía jurídica mexicana durante los últimos veinte años*, DTH-69

Sergio Visacovsky, *"Hasta la próxima crisis". Historia cíclica, virtudes genealógicas y la identidad de clase media entre los afectados por la debacle financiera en la Argentina (2001-2002)*, DTH-68

Rafael Rojas, *El debate de la Independencia. Opinión pública y guerra civil en México (1808-1830)*, DTH-67

Michael Sauter, *The Liminality of Man: Astronomy and the Birth of Anthropology in the Eighteenth Century*, DTH-66

Ventas

El **CIDE** es una institución de educación superior especializada particularmente en las disciplinas de Economía, Administración Pública, Estudios Internacionales, Estudios Políticos, Historia y Estudios Jurídicos. El Centro publica, como producto del ejercicio intelectual de sus investigadores, **libros, documentos de trabajo,** y **cuatro revistas especializadas:** *Gestión y Política Pública*, *Política y Gobierno*, *Economía Mexicana Nueva Época* e *Istor*.

Para adquirir cualquiera de estas publicaciones, le ofrecemos las siguientes opciones:

VENTAS DIRECTAS:	VENTAS EN LÍNEA:
Tel. Directo: 5081-4003 Tel: 5727-9800 Ext. 6094 y 6091 Fax: 5727 9800 Ext. 6314 Av. Constituyentes 1046, 1er piso, Col. Lomas Altas, Del. Álvaro Obregón, 11950, México, D.F.	Librería virtual: **www.e-cide.com** Dudas y comentarios: **publicaciones@cide.edu**

¡¡Colecciones completas!!

Adquiere los **CDs de las colecciones completas** de los documentos de trabajo de todas las divisiones académicas del CIDE: **Economía**, **Administración Pública**, **Estudios Internacionales**, **Estudios Políticos**, **Historia** y **Estudios Jurídicos.**

¡Nuevo! ¡¡Arma tu CD!!

Visita nuestra Librería Virtual www.e-cide.com y selecciona entre 10 y 20 documentos de trabajo. A partir de tu lista te enviaremos un CD con los documentos que elegiste.